



Helsinki  
Center  
of  
Economic  
Research

Discussion Papers

# Testing the New Keynesian Model on U.S. and Aggregate Euro Area Data

Mikael Juselius  
Swedish School of Economics, RUESG and HECER

Discussion Paper No. 131  
November 2006

ISSN 1795-0562

HECER – Helsinki Center of Economic Research, P.O. Box 17 (Arkadiankatu 7), FI-00014  
University of Helsinki, FINLAND, Tel +358-9-191-28780, Fax +358-9-191-28781,  
E-mail [info-hecer@helsinki.fi](mailto:info-hecer@helsinki.fi), Internet [www.hecer.fi](http://www.hecer.fi)

# Testing the New Keynesian Model on U.S. and Aggregate Euro Area Data\*

## Abstract

I apply the Johansen and Swensen (1999, 2004) method of testing exact rational expectations within the cointegrated VAR (Vector Auto-Regressive) model, to testing the New Keynesian Model (NKM). This method permits the testing of rational expectation systems, while allowing for non-stationary data. The NKM is tested on quarterly U.S. and aggregate Euro area time series data. I find that the restrictions implied by the core equations of the NKM are rejected regardless of sample periods and measures of real marginal costs. I also provide tentative explanations of the favorable results obtained by previous researches.

**JEL Classification:** C32, C52, E31, E52.

**Keywords:** New Keynesian Phillips curve, cointegration, vector autoregressive model.

Mikael Juselius

Department of Economics  
University of Helsinki  
P.O. Box 17 (Arkadiankatu 7)  
FI-00014 University of Helsinki  
FINLAND

e-mail: [mikael.juselius@helsinki.fi](mailto:mikael.juselius@helsinki.fi)

\* I gratefully acknowledge the financial support provided by the Research Unit on Economic Structures and Growth (RUESG) in the University of Helsinki. I also wish to thank Rune Stenbacka, Erkki Koskela, and Katarina Juselius for numerous valuable comments and suggestions. I bear sole responsibility for all remaining errors and omissions.

# 1 Introduction

The popularity of the New Keynesian model (NKM) in recent years has led to numerous empirical attempts to evaluate the performance of the model. A recent overview of this literature is provided by Henry and Pagan (2004). The typical study has focused on one of the two “core” equations of the NKM, usually the New Keynesian Phillips curve (NKPC) since it captures the relevant features of price stickiness. For instance Gali and Gertler (1999), and Gali et al. (2001) find strong evidence in favor of the Phillips curve, using single equation GMM (General Method of Moments). Sbordone (2002) also reports favorable results by a slightly different approach<sup>1</sup>, while Fuhrer (1997) obtains less favorable results using ML (Maximum Likelihood). More recent contributions include Matheron and Maury (2004), McAdam and Willman (2004) and Roberts (2005). The second equation, the expectational “IS” curve, has been investigated by Fuhrer (2000) and more recently in Kara and Nelson (2004) and Fuhrer and Rudebusch (2004).

Since the early and important contributions, a number of empirical issues have been raised. The use of single equation estimation procedures has been criticized on the grounds that empirical identification requires a system approach. Furthermore, due to such problems as weak instruments, GMM estimates are likely to be very imprecise. Thorough discussions on these issues can be found in Ma (2002), Mavroeidis (2004), Rudd and Whelan (2005a, 2005b)<sup>2</sup>. These difficulties has led authors such as Linde (2005) and Giordani (2004) to consider a full system approach. Another, largely neglected, issue is the apparent non-stationary behavior of the data. This problem is noted and discussed by Bardsen et al. (2004) among others. If data is non-stationary, we have an additional reason to view the previous results with caution.

The aim of this paper is to test the core equations of the NKM within a cointegrated VAR (Vector Auto-Regressive) model on quarterly U.S. and aggregate Euro area time series data. The sample periods are 1960:1-2005:2 and 1970:1-2003:4 for the U.S. and the aggregate Euro area data, respectively.

---

<sup>1</sup>Sbordone (2002) uses the method of testing present value models, proposed by Campbell and Shiller (1987). She assumes that the data is stationary, although the method also allows for special cases when the data is non-stationary. The Campbell and Shiller (1987) technique with non-stationary data has recently been employed to the NKPC by Demery and Duck (2003) and Tillmann (2005).

<sup>2</sup>However, see also Gali et al. (2005) and Sbordone (2005) for answers to some of this criticism.

The restrictions implied by the core equations of the NKM are tested by the method of testing exact linear rational expectations (RE), proposed by Johansen and Swensen (1999, 2004). This method permits the estimation of RE systems while allowing for non-stationary data. It should also be noted that the Johansen and Swensen method provides a direct and formal way of testing the NKM, as opposed to the more informal evaluations of much of the previous literature.

This paper is closely related to those of Fanelli (2005) and Barkbu and Batini (2005). Fanelli (2005) uses a three-step method, presented in Fanelli (2002), to formally test a version of the NKPC within a cointegrated VAR model on aggregate Euro area data<sup>3</sup>. He rejects the NKPC specification. Barkbu and Batini (2005) apply the Johansen and Swensen method to the NKPC on aggregate Euro area data. They obtain favorable results for the NKPC using a minimal information set. However, it is doubtful that the information set that Barkbu and Batini (2005) use captures the main features of the inflation process, as discussed by Bardsen et al. (2004). This paper differs from Fanelli (2005) and Barkbu and Batini (2005) in at least two ways. First, an extended information set is used that, combined with the Johansen and Swensen method, allows for testing both core equations of the NKM rather than only focusing on the NKPC. Second, the model is also tested on U.S. data.

The results suggest that the evidence in favor of the core equations of the NKM, the IS curve and the new Keynesian Phillips curve, is weak. The restrictions implied by the equations are rejected on both U.S. and aggregate Euro area data. Sensitivity analysis with respect to different sample periods and measures of marginal costs do not change the results. In addition, necessary conditions for the equations of the NKM in the non-stationary case are also provided and discussed. The necessary conditions imply cointegration between the key variables and are easily tested on data. The necessary conditions are rejected in most cases. Interestingly, the NKPC is not rejected on Euro area data when labor's share is used as a measure of marginal costs. It is precisely for this measure and data that favorable results on the NKPC have been reported by, for instance, Gali and Gertler (1999). However, this cannot be viewed as providing overwhelming support for the NKPC, since

---

<sup>3</sup>There are some drawbacks to his method. In particular, one can only test one RE equation and need a condition almost similar to strong exogeneity of the forcing variables in the estimations. Furthermore, Fanelli's parametrization of the NKPC does not allow for the standard condition  $\varphi_{41} + \varphi_{43} = 1$  in (4) below.

the overall restrictions of the equation are rejected. Methods that rely on less formal evaluations of the NKPC, such as the size and significance of the marginal costs and forward terms, run the risk of claiming success when only the necessary condition is met.

The next section introduces a baseline New Keynesian model. The data and information sets are discussed in section 3, while section 4 introduces the Johansen and Swensen method. The estimation results are presented in section 5 followed by a discussion in section 6. Section 7 concludes.

## 2 The New Keynesian model

This section introduces the baseline version of the New Keynesian model. The NKM belongs to a class of “miniature” dynamic stochastic general equilibrium (DSGE) models that are based on optimizing households and firms, rational expectations, and nominal price rigidities. The basic closed economy model has been derived through several different routes, for example by specifying different channels of price rigidities, as shown by Roberts (1995). Here we follow the standard approach, assuming Calvo pricing. Detailed derivations and discussions are provided by McCallum and Nelson (1999), Clarida et al. (1999), and Walsh (2003). The baseline model can be represented in terms of a two-equation model

$$y_t = E_t y_{t+1} - \varphi_{11}(i_t - E_t \Delta p_{t+1}) \quad (1)$$

$$\Delta p_t = \varphi_{21} E_t \Delta p_{t+1} + \varphi_{22} x_t \quad (2)$$

where  $y_t$  is real output,  $i_t$  is the nominal short-run interest rate,  $p_t$  is the price index,  $x_t$  is real marginal costs,  $E_t$  is the expectations operator conditional on the agent’s information set at time  $t$ , and  $\varphi_{ij} \geq 0$  for all  $i$  and  $j$  in equations (1)-(5). The first equation is a forward-looking “IS curve” that relates output to the real rate of interest, while the second is the New Keynesian Phillips curve that relates inflation to real marginal costs. In addition to this, a policy rule for the nominal interest rate is usually obtained by specifying a policy objective and solving under discretion or commitment. The coefficients,  $\varphi_{ij}$ , are functions of the structural parameter from the underlying theory.

Real marginal costs cannot be observed in practice and must, hence, be approximated by some measure. Under constant returns to scale and completely flexible nominal wages, real marginal cost are proportional to the deviation of output from its flexible price equilibrium, i.e.  $x_t = h(y_t - y_t^f)$ ,

where  $h > 0$  and  $y_t^f$  is the output level in the absence of price rigidities. In this case, it is common to rewrite (1) as

$$y_t - y_t^f = E_t(y_{t+1} - y_{t+1}^f) - \varphi_{11}(i_t - E_t\Delta p_{t+1}) + v_t$$

where,  $v_t = E_t y_{t+1}^f - y_t^f$ . Then, combining this equation with equation (2) and a policy rule for the interest rate produces a simple system of three equations in the endogenous variables  $x_t$ ,  $\Delta p_t$  and  $i_t$ . This is clearly convenient, especially from an econometric point of view. However, if the firms in the economy use Cobb-Douglas production technologies, the relevant measure of real marginal costs becomes labor's share of income, i.e.  $w_t n_t / y_t p_t$ , where  $w_t$  is wages and  $n_t$  is the number of employed. If this measure is used, rewriting (1) in terms of the output gap holds no advantage and, in principle, the process of the labor's share measure should be specified as well<sup>4</sup>. We return to this issue in section 3.1, where a simple solution to this problem is offered.

The purely theoretical equations (1) and (2) imply a jump behavior that is at odds with the observed behavior of both output and inflation (Fuhrer and Moore, 1995). This has led authors, for instance Fuhrer (2000) and Gali and Gertler (1999), to consider hybrid versions of (1) and (2) given by

$$y_t = \varphi_{32}E_t y_{t+1} - \varphi_{31}(i_t - E_t\Delta p_{t+1}) + \varphi_{33}y_{t-1} \quad (3)$$

$$\Delta p_t = \varphi_{41}E_t\Delta p_{t+1} + \varphi_{42}x_t + \varphi_{43}\Delta p_{t-1} \quad (4)$$

where the lagged terms are motivated, for example, by assuming habit persistence and rule of thumb pricing.

Finally, it should be noted that money is not absent from the New Keynesian model. Rather, it is redundant and determined by

$$m_t - p_t = \varphi_{51}y_t - \varphi_{52}i_t \quad (5)$$

where  $m_t$  is the nominal quantity of money. This 'unimportance of money' result is clearly testable empirically<sup>5</sup>.

The baseline model has been extended in several ways, for instance by incorporating labor market imperfections (Erceg et al., 2000) or by accounting for investments in capacity (Razin, 2005). Open economy issues have been investigated by several authors, for example Clarida et al. (2002), Gali and Monacelli (2002) and Svensson (2000). Such extensions are not considered

---

<sup>4</sup>A variant of this problem is discussed in Kara and Nelson (2004).

<sup>5</sup>See King (2002), for a discussion on the role for money.

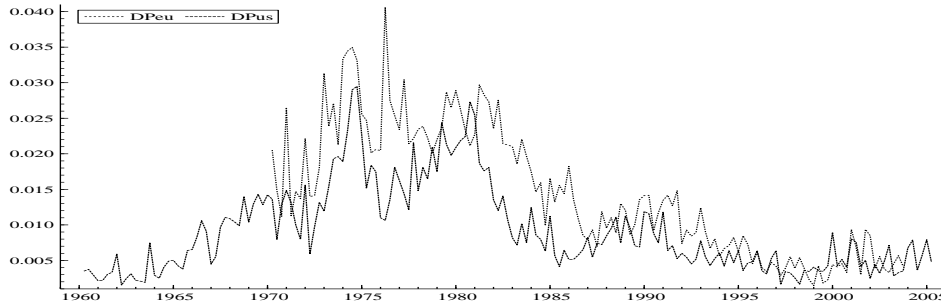


Figure 1: The aggregate Euro area and U.S. inflation rates.

in this paper. However, Clarida et al. (2002) notes that the open economy monetary policy problem is isomorphic to the closed economy problem and that the core equations of the NKM are almost identical under certain conditions. Hence, ignoring the open economy issues should not affect the results significantly.

### 3 Data and information

This section introduces the data and discusses potential information sets that can be used to evaluate the NKM. The data consists of quarterly U.S. and aggregate Euro area time series on the following variables (in logs): a price index,  $p_t$ , a nominal short-run interest rate,  $i_t$ , a real money aggregate,  $m_t$ , real output,  $y_t$ , potential output,  $y_t^n$ , and real aggregate wages,  $w_t$ . The Euro area data spans the years 1970:01-2003:04 while the U.S. data spans 1960:01-2005:02 (apart from a production function based measure of potential output which spans 1973:02-2003:04 and 1964:2-2005:02 respectively). Figure 1 plots the Euro area and U.S. inflation rates. Detailed descriptions of the data are provided in appendix A.

#### 3.1 Information sets

The minimal, and theory consistent, information set that can be used to test the baseline NKM is clearly  $\mathcal{I}_0 = \{\Delta p, i, y, x\}$ . This information set can be extended to include money,  $\mathcal{I}_1 = \{\Delta p, i, m, y, x\}$ , but if the theory is

correct this should be redundant. As this is a testable hypothesis it should not be simply assumed. Hence,  $\mathcal{I}_1$  will be the basic information set in this paper. It should be noted that most empirical studies on the NKPC work with a smaller information set  $\mathcal{I}_{npc} = \{\Delta p, x\}$  since the focus is on the Phillips curve<sup>6</sup>. Bardsen et al. (2004) have criticized the use of this kind of information set on the grounds that it is too small to account for the variation in the data, potentially leading to misspecified models. They show that this type of misspecification might explain the favorable results on the NKPC in the literature. When they extend the information set to include more variables, they find no support for the NKPC and that the results are consistent with a non-stationary inflation rate. Bardsen et al. take the view that inflation is primarily driven by labor market factors and extend their information set to include, wages, unemployment, productivity, taxes, and error correction mechanisms from existing studies. However, in some sense it is no longer possible to test the baseline NKPC when extending the information set in this way, since the additional information is not modelled within the NKM and rational expectations demand model consistency. The information set,  $\mathcal{I}_1$ , does not have this problem since it is consistent with the NKM and, as is confirmed below, is sufficient to ensure a well-specified model.

As the discussion in section 2 makes clear, different measures of real marginal costs have different implications for the core equations of the model. Which measure to use have been an issue of some debate previously (see Gali and Gertler (1999), Sbordone (2002) and Rudd and Whelan (2005)). This paper takes a pragmatic approach and uses several output gap measures and labor's share as proxies for real marginal costs. An elegant solution to the problems discussed in section 2 is to include the relevant measure in *unrestricted* form in the model. For example, if the preferred measure of marginal costs is the output gap, the information set,  $\mathcal{I}_{11} = \{\Delta p, i, m, y, y^n\}$ , is modelled. In this case the output gap is defined as the *restriction*,  $x_t = y_t - y_t^n$ , on the statistical model. If labor's share is used,  $\mathcal{I}_{12} = \{\Delta p, i, m, y, w\}$ , is modelled and labor's share is defined as the *restriction*,  $x_t = w_t - y_t$ . Note that this information should be in the agents' information sets since they can always deduce  $y_t^n$  or  $w_t$  from  $y_t$  and  $x_t$ . Hence, there is no particular reason to restrict the information at the outset. Figure 2 plots the EU output gap and

---

<sup>6</sup>Although, when GMM is used, the set of instruments usually contain other variables as well.



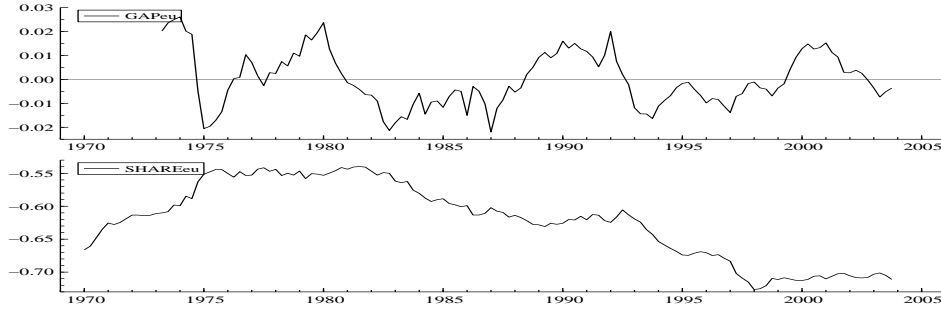


Figure 2: The production function based measure of the output gap and labor's share for the aggregate Euro area data.

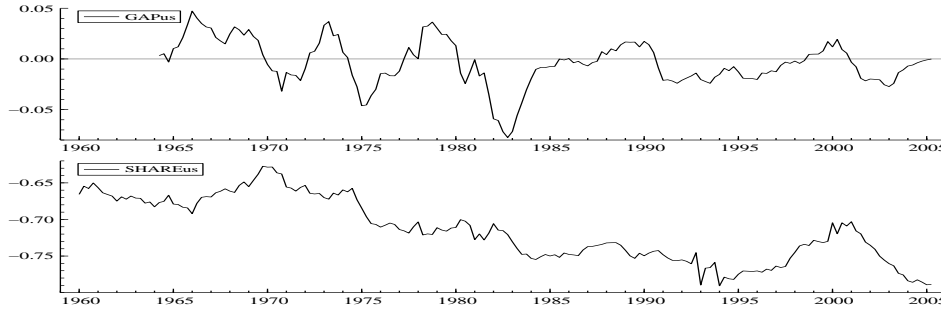


Figure 3: The production function based measure of the output gap and labor's share for the U.S. data.

labor's share and figure 3 plots the corresponding U.S. measures. It is clear from the figures that the two measures describe very different dynamics. In particular, the output gap measure appears to be in line with common views of the business cycle, while labor's share does not appear to capture cyclical variation to any large degree <sup>7</sup>.

As a final note, given the difficulties to obtain a reasonable measure for potential output, an alternative measure based on the Hodrick and Prescott (1997) filter was also used. There were no significant differences in the results<sup>8</sup>.

<sup>7</sup>See Rudd and Whelan, 2005 for a discussion of this point.

<sup>8</sup>Giorno et al. (1995) discusses the relative merits of different potential output measures.

## 4 Testing exact rational expectations within a cointegrated VAR model

This section describes the main results from Johansen and Swensen (2004) on testing rational expectations in a cointegrated VAR model when a linear trend is restricted to the cointegration space. The simpler case where there is no deterministic trend in the model is similar and described in Johansen and Swensen (1999). The exact restrictions implied by the NKM are presented at the end of the section.

The base line statistical model is the  $p$ -dimensional VAR model in error correction form

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu_0 + \mu_1 t + \Phi D_t + \varepsilon_t \quad (6)$$

where the vector process  $X_t$  is assumed to be at most  $I(1)$ ,  $\varepsilon_t \sim N_p(0, \Sigma)$ ,  $\Phi$  is a  $p \times m$  matrix, and  $D_t$  consist of the other deterministic components. Cointegration can be investigated as the hypothesis that the matrix  $\Pi$  is of reduced rank,  $r$ . If  $0 < r < p$  then at least some of the variables cointegrate and

$$\Pi = \alpha \beta'$$

where  $\alpha$  and  $\beta$  are two  $p \times r$  matrices of full column rank. Let the subscript  $\perp$  denote the orthogonal complement of a matrix. The deterministic trend is assumed to be restricted to the cointegration space, i.e.  $\alpha'_{\perp} \mu_1 = 0$ , since there would be quadratic trends in the data otherwise. Thus, we can write  $\mu_1 = \alpha \kappa_1$  where  $\kappa_1$  is an  $r$ -dimensional vector. These assumptions imply that (6) can be written as

$$\Delta X_t = \alpha \beta^{*'} X_{t-1}^* + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu_0 + \Phi D_t + \varepsilon_t \quad (7)$$

where  $\beta^* = (\beta', \kappa_1)'$  is a  $(p+1) \times r$  matrix and  $X_{t-1}^* = (X'_{t-1}, t)'$ .

Johansen and Swensen consider expectations of the form

$$\begin{aligned} E[c'_1 X_{t+1} \mid \Theta_t] + c'_0 X_t + c'_{-1} X_{t-1} + \dots + c'_{-k+1} X_{t-k+1} \\ + c_c + c_{\tau}(t+1) + c_{\phi} D_{t+1} = 0 \end{aligned} \quad (8)$$

where the  $p \times q$  ( $0 < q < r$ ) matrices  $c_i$  ( $i = -k+1, \dots, 1$ ) are known as well as the matrices  $c_{\tau}$  and  $c_{\phi}$ . The  $q$ -dimensional vector  $c_c$  can contain unknown

parameters. The expectational equation (8) can be reformulated so that it corresponds to (7) by

$$E[c'_1 \Delta X_{t+1} | \Theta_t] - d'_1 X_t + d'_{-1} \Delta X_{t-1} + \dots + d'_{-k+1} \Delta X_{t-k+2} + c_c + c_\tau(t+1) + c_\phi D_{t+1} = 0 \quad (9)$$

where  $d_{-i+1} = -\sum_{j=i-1}^{k-1} c_{-j}$ ,  $i = 0, \dots, k$ . Defining  $d_1^* = (d'_1, -c_\tau)'$ , the restrictions on the statistical model (7) implied by (9) are

$$\begin{aligned} \beta^* \alpha' c_1 &= d_1^* \\ \Gamma'_i c_1 &= -d_{-i} \\ \mu'_0 c_1 &= -c'_c \\ \Phi' c_1 &= -c'_\phi. \end{aligned} \quad (10)$$

The maximum likelihood under the restrictions is

$$L_{H,max}^{-2/T} = |\tilde{\Sigma}_{22}^*| |S_{11}^*| \prod_{i=1}^{r-q} (1 - \tilde{\lambda}_i^*) / |c'_1 c_1| |c'_{1\perp} c_{1\perp}| \quad (11)$$

where  $\tilde{\Sigma}_{22}^*$  is the likelihood of the marginal model,  $c'_1 \Delta X_t$ , and the remaining terms are the likelihood of the conditional model,  $c'_{1\perp} \Delta X_t$ . The product in (11) is taken to be 1 if  $q = r$ . The maximum likelihood of the unconstrained model (7) is

$$L_{max}^{-2/T} = |S_{00}^*| \prod_{i=1}^r (1 - \hat{\lambda}_i^*).$$

The LR test statistic, given as -2 times the log of the ratio between the restricted and the unrestricted likelihoods, is

$$\begin{aligned} -2\ln Q &= T \left( \ln |\tilde{\Sigma}_{22}^*| + \ln |S_{11}^*| + \sum_{i=1}^{r-q} \ln(1 - \tilde{\lambda}_i^*) \right) \\ &\quad - T \left( \ln |S_{00}^*| + \sum_{i=1}^r \ln(1 - \hat{\lambda}_i^*) + \ln(|c'_1 c_1| |c'_{1\perp} c_{1\perp}|) \right). \end{aligned}$$

The test statistic is asymptotically  $\chi^2$ -distributed with  $kpq + q(m+1)$  degrees of freedom. Although it is assumed that the  $c_i$  matrices are known, estimates of any unknown parameters,  $\varphi$ , in the  $c_i$  matrices can be obtained by numerical optimization, if the cointegrating relations can be expressed as smooth functions,  $\beta(\varphi)$ , of the parameters. In that case, the degrees of freedom turn out to be  $kpq + q(m+1) - w$ , where  $w$  is the number of additional unknown parameters. The core equations of the NKM satisfy this condition as is evident from the representations of  $d_1^*$  below.

## 4.1 Restrictions implied by the NKM

Let  $X_t = (\Delta p_t, i_t, m_t, y_t, y_t^n)'$  and  $k = 1$ . In terms of (8) the pure NKPC, equation (2), take the form

$$(-\varphi_{21}, 0, 0, 0, 0)E_t \begin{pmatrix} \Delta p_{t+1} \\ i_{t+1} \\ m_{t+1} \\ y_{t+1} \\ y_{t+1}^n \end{pmatrix} + (1, 0, 0, -\varphi_{22}, \varphi_{22}) \begin{pmatrix} \Delta p_t \\ i_t \\ m_t \\ y_t \\ y_t^n \end{pmatrix} = 0.$$

Hence,  $c_1 = (-\varphi_{21}, 0, 0, 0, 0)'$  and  $c_0 = (1, 0, 0, -\varphi_{22}, \varphi_{22})'$  which implies  $d_1 = (\varphi_{21} - 1, 0, 0, \varphi_{22}, -\varphi_{22})'$ . It is now straightforward to derive the restrictions on the parameters of (7) by using (10). Similarly, the pure IS curve in equation (1) can be expressed in terms of (8) by,  $c_1 = (-\varphi_{11}, 0, 0, -1, 0)'$  and  $c_0 = (0, \varphi_{11}, 0, 1, 0)'$ . The extensions to equations (3) and (4) are obvious, provided that  $k \geq 2$ , as are the consequences of using  $X_t = (\Delta p_t, i_t, m_t, y_t, w_t)'$  (the IS curve restrictions are the same, while the signs on  $\varphi_{22}$  are interchanged in  $c_0$  and  $d_1$  for the NKPC).

The simultaneous test of (1) and (2) can be performed by

$$c_1 = \begin{pmatrix} -\varphi_{11} & -\varphi_{21} \\ 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & 0 \end{pmatrix}, c_0 = \begin{pmatrix} 0 & 1 \\ \varphi_{11} & 0 \\ 0 & 0 \\ 1 & -\varphi_{22} \\ 0 & \varphi_{22} \end{pmatrix}$$

provided that  $r \geq 2$ . Similar extensions as above are again obvious.

## 5 Testing the NKM

In this section, the restrictions implied by the NKM are tested on Euro area and U.S. data. Initial modelling of the data is performed prior to testing the restrictions, since information about cointegration rank is a prerequisite in the Johansen and Swensen method. We begin by analysing aggregate Euro area data and then proceed with U.S. data.

$r$	$\mathcal{I}_{11}^{EU}, 1973:3-2003:4$				$\mathcal{I}_{12}^{EU}, 1970:1-2003:4$			
	$\lambda_i$	trace	trace95	p-value	$\lambda_i$	trace	trace95	p-value
0	0.41	158.42**	88.55	0.00	0.34	152.30**	88.55	0.00
1	0.27	95.83**	63.66	0.00	0.32	96.15**	63.66	0.00
2	0.25	57.76**	42.77	0.00	0.14	45.34*	42.77	0.03
3	0.14	23.43	25.73	0.10	0.12	25.50	25.73	0.05
4	0.04	5.18	12.48	0.58	0.07	9.08	12.48	0.18

Table 1: The rank test statistic (trace test) for the complete sample aggregate Euro area data. In the table,  $\lambda_i$  are the eigenvalues from the reduced rank regression (see Johansen, 1995). Trace95 are the 95%-quantiles of the trace distribution and (\*\*) denotes rejection at the 1% significance level and (\*) denotes rejection at the 5% significance level.

## 5.1 Aggregate Euro area data

This section reports the results of fitting the cointegrated VAR model (7) to the Euro area data with  $X_t^* = (\Delta p_t, i_t, m_t, y_t, y_t^n, t)'$  for the information set  $\mathcal{I}_{11}^{EU}$  and  $X_t^* = (\Delta p_t, i_t, m_t, y_t, w_t, t)'$  for  $\mathcal{I}_{12}^{EU}$ . The first model is referred to as the “gap model” and the second as the “share model” in what follows. Initial modelling suggested that  $k = 2$  is the suitable choice of lag length in both models and that linear trends should be included in the cointegration spaces.

Table 1 reports the rank test statistic of the models. The rank test statistic suggest that the rank should be set to three in both models, indicating three cointegration relations and two common stochastic trends. However, it can be seen from table 1 that  $r = 2$  was borderline accepted in the share model<sup>9</sup>. In fact, this seems to be the reasonable choice if one takes into account collateral evidence, such as the magnitude of the roots of the companion matrix, graphs of the implied CI relations under the different choices, and so on. Standard misspecification tests indicated some deviations from normality in both models<sup>10</sup>, as well as problems with auto correlation and ARCH in the share model. Stationarity and long-run exclusion were rejected in all variables in both models. The results of these tests are reported in

<sup>9</sup>A sensitivity analysis was conducted with respect to this choice but it did not change any of the results significantly. The results are available upon request from the author.

<sup>10</sup>This is mainly due to some large outliers in the turbulent seventies. These outliers can be accounted for by dummy variables, but doing so does not change the results.

r	$\mathcal{I}_{11}^{EU}, 1982:1-2003:4$				$\mathcal{I}_{12}^{EU}, 1982:1-2003:4$			
	$\lambda_i$	trace	trace95	p-value	$\lambda_i$	trace	trace95	p-value
0	0.41	133.54**	88.55	0.00	0.45	112.46**	88.55	0.00
1	0.27	88.67**	63.66	0.00	0.24	60.50	63.66	0.09
2	0.25	46.32*	42.77	0.02	0.19	37.04	42.77	0.17
3	0.14	21.03	25.73	0.18	0.11	18.58	25.73	0.31
4	0.04	7.17	12.48	0.34	0.09	8.40	12.48	0.23

Table 2: The rank test statistic (trace test) for the 1982:1-2003:2 aggregate Euro area data. In the table,  $\lambda_i$  are the eigenvalues from the reduced rank regression (see Johansen, 1995). Trace95 is the 95%-quantiles of the trace distribution and (\*\*) denotes rejection at the 1% significance level and (\*) denotes rejection at the 5% significance level.

appendix C. Finally, recursive tests for constant parameters were also performed on both models<sup>11</sup>. The results from these tests indicated two possible structural breaks, one in the early 1980's and one at around the middle of 1993. For this reason, separate analyses for the full sample and for the subsample 1982:1-2003:4 were conducted<sup>12</sup>. Similar breaks have been found, for example, by Batini (2002) and Barkbu and Batini (2005).

Table 2 reports the rank test statistic for the subsample 1982:1-2003:4. Again, the appropriate choice of rank seems to be three in the gap model, although  $r = 2$  was almost accepted (see footnote 9). For the share model the table indicates  $r = 1$ . However, similar arguments as previously, based on collateral evidence, would lead us to conclude that the rank should in fact be two (see footnote 9). There were no serious misspecification problems in either model, apart from some small deviations from normality. As before, stationarity was rejected in all variables in both models, but now long-run exclusion of  $w_t$  could not be rejected, with p-values 0.30 and 0.51 given  $r = 1$  and  $r = 2$ , respectively<sup>13</sup>. Finally, recursive tests of parameter stability

<sup>11</sup>The full description of these recursive test can be found in Juselius and Hansen (1995) and include two tests for the constancy of the  $\beta$ -vectors, a test for the constancy of the log-likelihood, a fluctuation test of the eigenvalues, among others. These results are available upon request.

<sup>12</sup>The samples 1970:1-1981:4 and 1993:2-2003:4 are not considered, since the subsamples are too small for effective estimations. However, the latter subsample will be commented on in appendix C.

<sup>13</sup>The trace test point unambiguously toward  $r = 2$  if  $w_t$  is excluded from the model.

were re-performed on the models. The tests did not show any parameter instability over the period. From an empirical point of view the first model is a slight favorite, since the labor's share measure does not seem to provide additional information apart from that which is contained in  $y_t$ . However, we will continue to use both measures in the rest of the paper.

Another issue that deserves comment is the fact that the real money supply seems to be needed in the information set. Long-run exclusion was rejected for this variable and removing it from the informations set considerably worsened the fit of each model. One of the results of the New Keynesian model is that money is determined by the other variables once a policy rule for interest rate is specified. Empirically, a necessary condition for the 'unimportance of money' result is that money has a unit vector in the  $\alpha$  matrix. This is formally tested in appendix C and the unit vector is rejected in all cases except for the 1982:1- share model. Thus, it would appear that money is important, at least when M3 is used as the money stock measure.

The results of testing the restrictions implied by the core equations of the NKM are reported in table 3. The details of the estimations are provided in appendix B. The single equation restrictions are first considered separately. As is clearly seen from table 3, almost all restrictions are strongly rejected. Furthermore, the coefficient estimates are clearly not sensible with respect to the NKM. For instance, in the cases of the NKPC, equations (2) and (4), the coefficients on the forward terms,  $\varphi_{21}$  and  $\varphi_{41}$  are above one and the coefficient on the forcing variable is small and negative *regardless of the measure used for marginal costs*. Also, for the IS curve, the coefficient on the real interest rate has the wrong sign, while the coefficients on the forward and backward terms can be considered sensible. In the few cases where the restrictions are not rejected, the coefficients are not sensible. The restrictions implied by the equations of the NKM are strongly rejected, if the coefficients are restricted to the unit interval (the results are available upon request).

Finally, the restrictions from both (1)-(2) and (3)-(4) where tested simultaneously on all periods and all information sets. These restrictions were strongly rejected in all cases, as should be expected, given the rejection of the single equation restrictions above.

These results imply that the evidence for the IS curve and the New Key-

---

Since additional information should not in principle reduce the CI rank, this provides an additional reason for maintaining  $r = 2$  despite the evidence for  $r = 1$  in table 2. It seems that the inclusion of  $w_t$  in the information set "muddles the water".

$\mathcal{I}$	$T$	Equ $i$	$\varphi_{i1}$	$\varphi_{i2}$	$\varphi_{i3}$	$-2\ln Q$	df	p-value
$\mathcal{I}_{11}^{EU}$	73:3-	1	-0.18	–	–	49.68	13	0.00
	“	2	1.11	-0.06	–	48.98	12	0.00
	“	3	-0.09	0.79	0.21	41.10	12	0.00
	“	4	1.79	-0.11	-0.66	30.44	11	0.01
	82:1-	1	-0.18	–	–	58.45	13	0.00
	“	2	1.16	-0.05	–	41.27	12	0.00
	“	3	-0.14	0.83	0.17	54.52	12	0.00
	“	4	1.96	-0.10	-0.87	20.86	11	0.04
$\mathcal{I}_{12}^{EU}$	70:1-	1	-0.33	–	–	55.83	13	0.00
	“	2	1.49	-0.05	–	55.72	12	0.00
	“	3	-0.20	0.81	0.19	48.73	12	0.00
	“	4	2.12	-0.04	-0.72	40.69	11	0.00
	82:1-	1	-0.22	–	–	44.10	13	0.00
	“	2	2.13	-0.10	–	14.28	12	<b>0.28</b>
	“	3	-0.17	0.83	0.17	40.24	12	0.00
	“	4	2.98	-0.11	-0.72	7.22	11	<b>0.78</b>

Table 3: Tests of the restrictions implied by the core equations of the NKM (1)-(4) on the Euro area data. The column “Equ  $i$ ” indicates that the restrictions implied by equation ( $i$ ) is being tested and  $\varphi_{ij}$  are the corresponding estimates. In equation (3) we have the additional restriction  $\varphi_{32} + \varphi_{33} = 1$  (hence, we have 12 degrees of freedom).

nesian Phillips curve on aggregate Euro area data must be considered weak. The results of testing the NKPC are similar to those of Fanelli (2005) in this respect. In section 6 we discuss some reasons for this failure of the model.

## 5.2 U.S. data

Initial modelling of the two information sets for U.S. data suggested  $k = 3$  and that a linear trend should be included in the cointegration space in both models. Table 4 reports the rank test statistic of the model. The rank test statistic suggests that the rank should be set to two in the first model. However,  $r = 1$  is almost accepted in the share model and, moreover, there is uncertainty between the choices  $r = 2$  or  $r = 3$  as well. It appears that the inclusion of  $w_t$  in the information set again “muddles the water”. The choice  $r = 3$  is not reasonable and can be disregarded if one takes



$\mathcal{I}_{11}^{US}, 1964:2-2005:2$					$\mathcal{I}_{12}^{US}, 1960:1-2005:2$			
r	$\lambda_i$	trace	trace95	p-value	$\lambda_i$	trace	trace95	p-value
0	0.24	114.22**	88.55	0.00	0.22	108.36**	88.55	0.000
1	0.19	70.06*	63.66	0.01	0.11	64.04*	63.66	0.046
2	0.13	36.86	42.77	0.18	0.09	42.66	42.77	0.051
3	0.06	14.55	25.73	0.62	0.07	24.40	25.73	0.074
4	0.03	5.15	12.48	0.58	0.06	10.18	12.48	0.121

Table 4: The rank test statistic (trace test) for the complete sample U.S. data. In the table,  $\lambda_i$  are the eigenvalues from the reduced rank regression (see Johansen, 1995). Trace95 is the 95%-quantiles of the trace distribution and (\*\*) denotes rejection at the 1% significance level and (\*) denotes rejection at the 5% significance level.

into account similar collateral evidence as before. Below, only the results for  $r = 2$  are reported (see footnote 9). Standard misspecification tests indicated deviations from normality, due to some very large outliers, and problems with ARCH, stemming from the short-term interest rate series, in both models. None of the variables in either model were found to be stationary or long-run excludable. Finally, recursive tests for constant parameters were also performed. Both models showed evidence of a structural break at around 1979, marking the beginning of the Volcker-Greenspan era. Similar structural breaks have been found in the empirical literature, for example in Roberts (2005) and Romer and Romer (2004). However, by some experimenting, it can be seen that the ARCH problems in the short-run interest rate series do not disappear before 1982, so it seems a good idea to split the sample at that point. Roberts (2005) considers a similar split by leaving out the years 79-83 corresponding to the Volcker disinflation era. A sensitivity analysis with respect to this choice was conducted, but it did not change the main results. Thus, separate analyses of the full sample and of the subsample 1982:1-2005:2 were conducted<sup>14</sup>. Additionally, there were also some evidence of a structural break around 1993 in the share model.

Table 5 reports the rank test statistic for the sample 1982:1-2005:2. The appropriate choice of rank is two in both models. There was no serious misspecification in the gap model, apart from some small deviations from normality, while there were evidence of small problems with ARCH, auto-

<sup>14</sup>It would also be possible to conduct a separate analysis for the period 1960:1-1981:4.

r	$\mathcal{I}_{11}^{US}, 1982:1-2005:2$				$\mathcal{I}_{12}^{US}, 1982:1-2005:2$			
	$\lambda_i$	trace	trace95	p-value	$\lambda_i$	trace	trace95	p-value
0	0.53	147.41**	88.55	0.00	0.51	137.34**	88.55	0.00
1	0.35	78.80**	63.66	0.00	0.36	73.83**	63.66	0.00
2	0.21	39.61	42.77	0.10	0.17	34.01	42.77	0.29
3	0.10	18.71	25.73	0.31	0.15	17.44	25.73	0.39
4	0.09	8.92	12.48	0.19	0.03	2.80	12.48	0.88

Table 5: The rank test statistic (trace test) for the subsample, 1982:1-2005:2, U.S. data. In the table,  $\lambda_i$  are the eigenvalues from the reduced rank regression (see Johansen, 1995). Trace95 is the 95%-quantiles of the trace distribution and (\*\*) denotes rejection at the 1% significance level and (\*) denotes rejection at the 5% significance level.

correlations, and deviations from normality in the second model. Stationarity was rejected in all variables in both models and again the long-run exclusion of  $w_t$  could not be rejected (p-value 0.06) in the second model. Finally, recursive tests for parameter stability was re-performed for the models. The tests did not show any serious parameter instability over the period in the gap model, while there were still some evidence of a break around 1993 in the share model.

The results of testing the restrictions implied by the core equations of the NKM on the U.S. data are reported in table 6. Almost all restrictions are rejected, as can be seen from table 6. Furthermore, the coefficient estimates are very similar to those of the aggregate Euro area data. Finally, the restrictions from both (1)-(2) and (3)-(4) were tested simultaneously on both periods and both models. These restrictions were strongly rejected in all cases. Hence, the evidence in favor of the NKM on U.S data must also be considered weak.

## 6 Explaining the results

The reasons for the empirical failure of the NKM are investigated in this section. We begin by discussing a particular necessary condition for the NKM. This condition is then tested and interpreted in light of previous findings in the literature. The estimated coefficients in tables 3 and 6 are also given an interpretation.

$\mathcal{I}$	$T$	Equ $i$	$\varphi_{i1}$	$\varphi_{i2}$	$\varphi_{i3}$	$-2\ln Q$	df	p-value
$\mathcal{I}_{11}^{US}$	64:2-	1	-0.40	—	—	86.41	18	0.00
	“	2	1.13	-0.03	—	49.13	17	0.00
	“	3	-0.28	0.80	0.20	75.01	17	0.00
	“	4	1.52	-0.04	-0.54	24.43	16	<b>0.08</b>
	82:1-	1	0.06	—	—	72.63	18	0.00
	“	2	1.60	-0.02	—	50.71	17	0.00
	“	3	-0.27	0.80	0.20	75.53	17	0.00
	“	4	1.98	-0.07	-0.75	39.02	16	0.00
$\mathcal{I}_{12}^{US}$	60:1-	1	-0.43	—	—	67.84	18	0.00
	“	2	1.16	-0.01	—	38.16	17	0.00
	“	3	-0.29	0.80	0.20	55.44	17	0.00
	“	4	1.62	-0.01	-0.45	23.43	16	<b>0.10</b>
	82:1-	1	0.06	—	—	65.35	18	0.00
	“	2	1.63	-0.01	—	55.62	17	0.00
	“	3	0.01	0.71	0.29	48.41	17	0.00
	“	4	2.25	-0.01	-0.54	49.43	16	0.00

Table 6: Tests of the restrictions implied by the core equations of the NKM (1)-(4) on the U.S. data. The column “Equ  $i$ ” indicates that the restrictions implied by equation ( $i$ ) are being tested and  $\varphi_{ij}$  are the corresponding estimates. In equation (3) we have the additional restriction  $\varphi_{32} + \varphi_{33} = 1$  (hence, we have 17 degrees of freedom).

## 6.1 A necessary condition for the NKM under $I(1)$ data

It is easy to provide a necessary condition for the NKM provided that the data is non-stationary and well described by model (7), as the analysis in section 5 suggests. In this case inflation must be cointegrated with the measure of marginal costs and output must be cointegrated with the real rate of interest<sup>15</sup>. This can be seen directly, by observing that if the first restriction

$$\beta^* \alpha' c_1 = d_1^* \quad (12)$$

in (10) holds, then  $d_1^* \in sp(\beta^*)$ . That  $d_1^* \in sp(\beta^*)$  is only a necessary condition is clear, since (10) includes several other restrictions. The advantage of the necessary condition  $d_1^* \in sp(\beta^*)$  is that it is very easy to verify on data.

<sup>15</sup>A similar necessary condition on cointegration for present value models is discussed by Campbell and Shiller (1987).

$\mathcal{I}^{EU}$		$\hat{\beta}_{\Delta p}$	$\hat{\beta}_i$	$\hat{\beta}_m$	$\hat{\beta}_y$	$\hat{\beta}_{y^w/w}$	$\hat{\beta}_t$	p-value
$\mathcal{I}_{11}^{EU}$	$d_{11}^*$	-1	1	0	0	0	0	0.00
	$d_{21}^*$	1	0	0	7.78	-7.78	0	0.01
	$d_{31}^*$	-5.17	5.17	0	1	0	0	0.00
	$d_{41}^*$	0	0	0	1	-1	0	0.02
$\mathcal{I}_{12}^{EU}$	$d_{11}^*$	-1	1	0	0	0	0	0.00
	$d_{21}^*$	1	0	0	0.08	-0.08	0	<b>0.37</b>
	$d_{31}^*$	-26.42	26.42	0	1	0	0	0.03
	$d_{41}^*$	0	0	0	-1	1	0	0.00
$\mathcal{I}_{11}^{US}$	$d_{11}^*$	-1	1	0	0	0	0	0.00
	$d_{21}^*$	1	0	0	-0.18	0.18	0	<b>0.33</b>
	$d_{31}^*$	-41.10	41.10	0	1	0	0	0.00
	$d_{41}^*$	0	0	0	1	-1	0	<b>0.10</b>
$\mathcal{I}_{12}^{US}$	$d_{11}^*$	-1	1	0	0	0	0	0.00
	$d_{21}^*$	1	0	0	0.03	-0.03	0	0.00
	$d_{31}^*$	-46.70	46.70	0	1	0	0	0.01
	$d_{41}^*$	0	0	0	-1	1	0	0.00

Table 7: Tests of the necessary conditions of the NKM. Estimated  $\beta$  coefficients are denoted by  $\hat{\beta}_x$ , where  $x$  indicates the variable. The hypotheses in  $d_{41}^*$  are derived under the additional restriction  $\varphi_{41} + \varphi_{43} = 1$ .

For each of equations (1)-(4),  $d_1^*$  will take an explicit form (the  $d_1^*$  corresponding to equation  $i$  is denoted by  $d_{i1}^*$ ). Thus, if  $x_t = y_t - y_t^n$  we get

$$\begin{aligned}
d_{11}^* &= (-\varphi_{11}, \varphi_{11}, 0, 0, 0, 0)' \\
d_{21}^* &= (1 - \varphi_{21}, 0, 0, -\varphi_{22}, \varphi_{22}, 0)' \\
d_{31}^* &= (-\varphi_{31}, \varphi_{31}, 0, 1 - \varphi_{32} - \varphi_{33}, 0, 0)' \\
d_{41}^* &= (1 - \varphi_{41} - \varphi_{43}, 0, 0, -\varphi_{42}, \varphi_{42}, 0)'
\end{aligned}$$

and if  $x_t = w_t - y_t$ , the signs on the coefficients  $\varphi_{22}$  and  $\varphi_{42}$  are interchanged. Assumed  $\varphi_{ij} > 0$  for all  $i$  and  $j$ , it can be seen that  $d_{11}^*$  is nested in  $d_{31}^*$  by the restriction  $\varphi_{32} + \varphi_{33} = 1$ , and that  $d_{21}^*$  and  $d_{41}^*$  are similar. Note also the theoretically interesting cases  $\varphi_{21} \neq 1$  and  $\varphi_{41} + \varphi_{43} = 1$ . Table 7 reports the results of testing if these relations are in the estimated cointegration space. Attention is restricted to the subsample starting in 1982:1. It can be seen from the table, that the necessary condition of the “IS” curve, the

rows of  $d_{11}^*$  and  $d_{31}^*$ , is rejected in all cases. Hence, it is not surprising that the corresponding hypotheses in tables 3 and 6 were rejected. Table 7 also reveals some interesting facts about the NKPC ( $d_{21}^*$  and  $d_{41}^*$ ). The necessary condition of the NKPC on the Euro area data is rejected if we use the output gap as a measure of marginal cost. Furthermore, it can easily be seen that the signs of the estimates are wrong in this case. However, if labor's share is used as a measure instead, the necessary condition holds and the signs on the coefficients are correct and of reasonable magnitude. This, then, provides a possible explanation for the success of the labor's share measure in the previous literature. In essence, since the data has been assumed to be stationary even though this is clearly rejected if tested, what has been estimated is the necessary condition on cointegration implied by the NKPC. However, this finding is not sufficient to conclude that the NKPC is a good description of inflation. Methods that rely on less formal tests of the NKPC, and more generally the NKM, run the risk of claiming success when only the necessary condition is met, in particular when data is non-stationary. Generally, cointegration between the key variables of any expectational equation of the form (8), is a necessary condition when the data is  $I(1)$ . Hence, investigating cointegration between the variables in the system provides valuable information for future theoretical developments. This, potentially interesting avenue, is not explored further in the present paper.

The opposite finding holds on U.S. data. The necessary condition is not rejected when the output gap measure is used, but rejected when the labor's share measure is used. This result might account for the poor performance of the labor's share based NKPC on U.S. data that has been previously observed.

Finally, note that  $d_{41}^*$  in table 7 essentially tests if the output gap or labor's share are stationary. The stationarity of the measures is rejected in most cases, with the exception of the U.S. output gap. However, output gap measures should be structurally stationary by construction and it is puzzling that stationarity is rejected the European output gap. It appears that the short length of the subsample and the persistence of the European business cycle creates a near unit-root problem in the output gap series<sup>16</sup>.

---

<sup>16</sup>The unit-root in the output gap variable is rejected if the full sample, 1973:2-2003:4, is investigated. This illustrates the fact that we have long and persistent business cycles.

## 6.2 Coefficient estimates and solution of the RE system

The fact that neither  $x_t$  and  $\Delta p_t$  nor  $y_t$  and  $r_t = i_t - \Delta p_t$  are cointegrated in most cases, may account for the implausible and strange coefficients in tables 3 and 6. For instance, note that the coefficients on  $x_t$  in the NKPC are consistently small compared to the coefficients on  $\Delta p_t$ . If the coefficients on  $x_t$  are not statistically different from zero, it seems reasonable that the coefficients capture the unit root behavior of inflation, rather than being meaningful in terms of the NKPC. Such results were found by Bardsen et al. (2004) on aggregate Euro area data. The stability of the RE system can be investigated by the method proposed by Blanchard and Kahn (1980). To this end, (1)-(2) and (3)-(4), with  $x_t = y_t - y_t^n$ , are written in the form

$$\begin{pmatrix} X_{t+1} \\ E_t P_{t+1} \end{pmatrix} = A \begin{pmatrix} X_t \\ P_t \end{pmatrix} + \gamma Z_t \quad (13)$$

where  $\gamma Z_t$  collects the exogenous variables  $y_t^n$  and  $i_t$ <sup>17</sup>. In terms of (13), (1) and (2) are represented by  $X_t = \emptyset$ ,  $P_t = (y_t, \Delta p_t)'$ , and

$$A = \begin{pmatrix} 1 - \frac{\varphi_{11}\varphi_{22}}{\varphi_{21}} & -\frac{\varphi_{11}}{\varphi_{21}} \\ \frac{-\varphi_{22}}{\varphi_{21}} & \frac{1}{\varphi_{21}} \end{pmatrix}. \quad (14)$$

Likewise, for equations (3) and (4) we have  $X_t = (y_{t-1}, \Delta p_{t-1})'$ ,  $P_t = (y_t, \Delta p_t)'$ , and

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \frac{-\varphi_{33}}{\varphi_{32}} & \frac{\varphi_{31}\varphi_{43}}{\varphi_{32}\varphi_{41}} & \frac{1}{\varphi_{32}} + \frac{\varphi_{31}\varphi_{42}}{\varphi_{32}\varphi_{41}} & \frac{-\varphi_{31}}{\varphi_{32}\varphi_{41}} \\ 0 & \frac{-\varphi_{43}}{\varphi_{41}} & \frac{-\varphi_{42}}{\varphi_{41}} & \frac{1}{\varphi_{41}} \end{pmatrix}. \quad (15)$$

The corresponding  $A$  matrices when  $x_t = w_t - y_t$  are similar apart from some changes in the signs. Using the values from tables 3 and 6, the roots of (14) and (15) can be calculated. For all cases in the tables, one root is very close to unity, while the remaining roots are within the unit circle. In the different gap models the roots cluster around 1.02 and for the share models around 0.99. This suggest that there is no unique stable forward solution to the system and, since it is unlikely that we could reject the unit root statistically, that the

---

<sup>17</sup>Alternatively, both  $y_t^n$  and  $i_t$  could be treated as predetermined with roots less or equal to one in absolute value. Doing so does not add anything to the analysis.

solution is non-stationary. Hence, the proper interpretation of the estimates in tables 3 and 6 is that they confirm that the data is non-stationary and that the necessary conditions do not hold.

## 7 Conclusions

This paper applies the Johansen and Swensen (1999, 2004) method for testing linear rational expectations models, to testing the New Keynesian Model on U.S. and aggregate Euro area data. The tests were conducted on both the individual equations separately and on system as a whole. The NKM was rejected on both U.S. and aggregate Euro area data. Several sensitivity analyses with respect to the choice of measures, sample periods, etc. were also preformed but they did not change the results. Hence, the evidence for the NKM must be considered weak.

Some potential reasons for the empirical shortcomings of the model were also discussed. Much of the previous literature has assumed stationarity on behalf of the key variables in the NKM. However, this is empirically implausible, as shown in this paper among others. When non-stationarity is allowed, the equations of the NKM do not satisfy, in most cases, a particular necessary condition, namely that the key variables must be cointegrated. Interestingly, the necessary condition is satisfied when labor's share is used as a measure of marginal costs on Euro area data. This might explain the success of the NKPC previously reported for this measure and data. In essence, what has previously been estimated is the necessary condition, i.e. a cointegration relationship. This has then been interpreted as evidence in favor of the NKPC, although a formal test of this hypothesis is rejected. The necessary conditions are not satisfied on U.S. data, which accounts for the previously reported poorer performance of the model on the U.S. data.

The results also suggest a potential way forward. Cointegration between the key variables, is necessary condition of any linear rational expectation hypothesis, when the data is non-stationary. Thus, any exploratory investigation on cointegration between the variables within a economically meaningful information set clearly provides valuable information on potential extensions of the theoretical models.

## A Data definitions and sources

This appendix provides the precise definitions and sources of the data that was used in the analysis. All data is available from the sources below (membership required for the AWM), or upon request from the author.

### A.1 Aggregate Euro area data

The main data source for the European data is the Area Wide Model (AWM) dataset, available from the Euro Area Business Cycle Network (EABCN, [www.eabcn.org](http://www.eabcn.org), see Fagan et al., 2001). Additional data was obtained from OECD databases. The data spans the years 1970:1-2003:4, with the notable exception of the production function based potential output series which spans 1973:2-2003:4.

$p_t$  = (log of) GDP deflator, base year 1995 (AWM series YED). A sensitivity analysis was conducted by using the CPI index (OECD, economic outlook) but it did not change the results significantly.

$r_t$  = Short-run interest rate (AWM series STN).

$m_t$  = (log of) Real EMU monetary aggregate M3 in millions of EUR. The nominal series was obtained from OECD, main economic indicators, and deflated by the index used for  $p_t$ .

$y_t$  = (log of) Real GDP (AWM series YER).

$y_t^n$  = (log of) Potential real output. The main measure used was the production function based measure from AWM series YET. This measure was available from 1973:2-2003:4. Sensitivity analysis was conducted with Hodric-Prescott filtered real GDP (using scale parameters 400, 1600).

$w_t$  = (log of) Total real compensation to employees (AWM series WIN deflated by  $p_t$ ).

It should be pointed out that the transformation,  $x_t = w_t - y_t$ , is identical to the labor's share measure used in Clarida et al. (1999), apart from scaling.



## A.2 U.S. data

The main source for the U.S. data is the OECD database ([www.oecd.org](http://www.oecd.org)). The data spans the years 1960:1-2005:2, with the notable exception of the production function based potential output series which spans 1964:2-2005:2.

$p_t$  = (log of) GDP deflator, base year 2000. A sensitivity analysis was conducted by using the CPI index but it did not change the results significantly. Both series can be found in the OECD economic outlook database.

$r_t$  = 3 month LIBOR, obtained from the OECD economic outlook database.

$m_t$  = (log of) Real money stock M2 in millions of US dollars (OECD, economic outlook). Deflated by  $p_t$ .

$y_t$  = (log of) Real GDP (OECD, economic outlook).

$y_t^n$  = (log of) Potential real output. The main measure used in the analysis was the production function based measure (available from OECD, economic outlook). This measure was available from 1964:2-2005:2. Sensitivity analysis was conducted with Hodric-Prescott filtered GDP (using scale parameters 400, 1600).

$w_t$  = (log of) Total real wages and salaries obtained from OECD, economic outlook (deflated by  $p_t$ ).

The transformation,  $w_t - y_t$ , corresponds very closely to the labor's share measure published by the Bureau of Labors Statistics (BLS, [www.bls.gov](http://www.bls.gov)). The main difference comes from the use of GDP instead of non-farm business sector output. A sensitivity analysis was conducted with respect to the output measure without altering the results.

## B Optimization

This appendix describes the methods used to obtain the coefficient estimates of the unknown parameters in the  $c_i$  matrices of section 4. As noted by Johansen and Swensen (1999), as long as the functions of the parameters are smooth, numerical optimization techniques can be applied to maximize the likelihood function. To this end both grid search and the quasi Newton

Test for stationarity								
$\mathcal{I}$	$T$	$r$	$\Delta p_t$	$i_t$	$m_t$	$y_t$	$y_t^n$	$w_t$
$\mathcal{I}_{11}^{EU}$	73:2-	3	0.00**	0.00**	0.00**	0.00**	0.00**	—
	82:1-	3	0.00**	0.00**	0.00**	0.00**	0.00**	—
$\mathcal{I}_{12}^{EU}$	70:1-	3	0.01**	0.00**	0.00**	0.01**	—	0.02*
	82:1-	2	0.01*	0.00**	0.00**	0.00**	—	0.01**
$\mathcal{I}_{11}^{US}$	64:2-	2	0.00**	0.00**	0.00**	0.00**	0.00**	—
	82:1-	2	0.00**	0.00**	0.00**	0.00**	0.00**	—
$\mathcal{I}_{12}^{US}$	60:1-	2	0.06	0.01*	0.00**	0.00**	—	0.00**
	82:1-	2	0.00**	0.00**	0.00**	0.00**	—	0.00**

Table 8: Test for stationarity. The table reports the p-values of the hypothesis. (\*) and (\*\*) indicates rejection at the 5% and 1% significance levels respectively.

optimization algorithm by Broyden-Fletcher-Goldfarb-Shanno (BFGS) were used.

In some of the cases there were several local maxima, in which case a grid search over reasonable starting values was conducted. The reported parameters correspond to the maximum (in all cases, the other local maxima produced very low values of the likelihood and very extreme values of the parameters).

Restricting the parameters to the unit interval was conducted by setting  $\varphi_{ij} = \frac{1}{1+|\mathcal{V}_{ij}|}$  and maximizing over  $\mathcal{V}_{ij}$ , and by grid search over the unit intervals. The hypotheses were strongly rejected in all cases.

## C Miscellaneous results

Various results that are of interest, but strictly not needed in the main text, are reported in this appendix. Table 8 reports the tests for stationarity<sup>18</sup>.

It was claimed in the text that a test for a unit vector in the  $\alpha$  matrix was rejected in the money equation. The results from testing this hypothesis on the subsample 1982:1- produced p-values 0.001 and 0.20 for the EU gap and share models respectively. Similarly, we get p-values 0.001 and 0.000 for

<sup>18</sup>The tests for trend stationarity were similar, apart from a few cases in the longer sample, where weak evidence for trend stationarity was found.

$\mathcal{I}$	Equ $i$	$\varphi_{i1}$	$\varphi_{i2}$	$\varphi_{i3}$	$-2\ln Q$	df	p-value
$\mathcal{I}_{11}^{EU}$	1	0.29	–	–	24.19	13	0.03*
$\mathcal{I}_{11}^{EU}$	3	0.12	0.70	0.30	15.59	12	0.21
$\mathcal{I}_{12}^{EU}$	1	0.18	–	–	33.28	13	0.00**
$\mathcal{I}_{12}^{EU}$	3	0.06	0.69	0.31	24.26	12	0.02*

Table 9: Tests of the restrictions implied by the equations (1) and (3) on the Euro area data subsample 1993:3-2003:4. The column “Equ  $i$ ” indicates that the restrictions implied by equation ( $i$ ) is being tested and  $\varphi_{ij}$  are the corresponding estimates. In equation (3) we have the additional restriction  $\varphi_{32} + \varphi_{33} = 1$  (hence, we have 12 degrees of freedom).

the US gap and share models respectively. The results from the full sample tests were similar.

Finally, table 9 provides the results from testing the optimizing IS curve on the Euro area data subsample 1993:3-2003:4. These results should be viewed with great caution since only 43 observations are used in the estimations. Nevertheless, the results point to the possibility of a structural break whereafter output evolves according to (1) or (3), at least when the output gap is used. The coefficient  $\varphi_{11}$ , and to some extent  $\varphi_{31}$ , depending on the assumptions used to derive the hybrid version, is the inverse of a preference parameter  $\sigma$ , where  $\sigma$  stems from an utility function of the form  $u(c, \cdot) = \frac{c^{1-\sigma}}{1-\sigma} + \dots$ . The two first equations in table 9 imply the estimates  $\sigma_{11} = 3.45$  and  $\sigma_{31} = 8.33$ , which are highly plausible. Furthermore, the weight to the forward variable is approximately 0.70 compared to 0.30 for the backward variable, in line with the beliefs of most researchers. However, the NKPC was rejected on this same sample, as were both the IS curve and the NKPC on the U.S. data.

## References

- Bardsen, G., Jansen, E. and Nymoen, R. (2004), ‘Econometric Evaluation of the New Keynesian Phillips Curve’, *Oxford Bulletin of Economics and Statistics* **66**, 671–686. Supplement.
- Barkbu, B. and Batini, N. (2005), The New-Keynesian Phillips Curve When Inflation is Non-Stationary, PhD thesis, European University Institute. In PhD thesis "Essays on Prices" by Bergljot Barkbu (chapter 3).
- Batini, N. (2002), Euro Area Inflation Persistence. ECB Working Paper No. 201.
- Blanchard, J. O. and Kahn, C. M. (1980), ‘The Solution of Linear Difference Models under Rational Expectations’, *Econometrica* **48**(5), 1305–1312.
- Campbell, J. Y. and Shiller, R. J. (1987), ‘Cointegration and Tests of Present Value Models’, *The Journal of Political Economy* **95**(5), 1062–1088.
- Clarida, R., Gali, J. and Gertler, M. (1999), ‘The Science of Monetary Policy: A New Keynesian Perspective’, *Journal of Economic Literature* **37**(4), 1661–1707.
- Clarida, R., Gali, J. and Gertler, M. (2002), ‘A simple framework for international monetary policy analysis’, *Journal of Monetary Economics* **49**, 879–904.
- Demery, D. and Duck, N. W. (2003), Cointegration-Based Tests of the New Keynesian Model of Inflation. University of Bristol Discussion Paper No. 541.
- Erceg, C. J., Henderson, D. W. and Levin, A. T. (2000), ‘Optimal monetary policy with staggered wage and price contracts’, *Journal of Monetary Economics* **46**, 281–313.
- Fagan, G., Henry, J. and Mestre, R. (2001), An Area-Wide Model (AWM) for the Euro Area. ECB Working Paper No. 42.
- Fanelli, L. (2002), ‘A new approach for estimating and testing the linear quadratic adjustment cost model under rational expectations and I(1) variables’, *Journal of Economic Dynamics & Control* **26**, 117–139.

- Fanelli, L. (2005), Testing the New Keynesian Phillips Curve through Vector Autoregressive Models: Results from the Euro Area. Universita Ca Foscari, Venezia, January 24-25, <http://www.cide.info/conf/congress.php>.
- Fuhrer, J. C. (1997), 'The (Un)Importance of Forward-Looking Behavior in Price Specifications', *Journal of Money, Credit and Banking* **29**(3), 338–350.
- Fuhrer, J. C. (2000), 'Habit Formation in Consumption and Its Implications for Monetary-Policy Models', *The American Economic Review* **90**(3), 367–390.
- Fuhrer, J. C. and Moore, G. (1995), 'Inflation Persistence', *The Quarterly Journal of Economics* **110**(1), 127–159.
- Fuhrer, J. C. and Rudebusch, G. D. (2004), 'Estimating the Euler equation for output', *Journal of Monetary Economics* **51**, 1133–1153.
- Gali, J. and Gertler, M. (1999), 'Inflation dynamics: A structural econometric analysis', *Journal of Monetary Economics* **44**, 195–222.
- Gali, J., Gertler, M. and Lopez-Salido, D. J. (2001), 'European inflation dynamics', *European Economic Review* **45**, 1237–1270.
- Gali, J., Gertler, M. and Lopez-Salido, D. J. (2005), 'Robustness of the estimates of the hybrid New Keynesian Phillips Curve', *Journal of Monetary Economics* **52**(6), 1107–1118.
- Gali, J. and Monacelli, T. (2002), Monetary Policy and Exchange Rate Volatility in a Small Open Economy. CEPR Discussion Papers 3346.
- Giordani, P. (2004), 'Evaluating New-Keynesian Models of Small Open Economy', *Oxford Bulletin of Economics and Statistics* **66**, 713–733. Supplement.
- Giorno, C., Richardson, P., Roseveare, D. and van den Noord, P. (1995), Estimating Potential Output, Output Gaps and Structural Budget Balances. OECD Economics Department Working Papers, No. 152.
- Henry, B. S. G. and Pagan, A. R. (2004), 'The Econometrics of the New Keynesian Policy Model: Introduction', *Oxford Bulletin of Economics and Statistics* **66**, 581–607. Supplement.

- Hodrick, R. J. and Prescott, E. C. (1997), 'Postwar U.S. Business Cycles: An Empirical Investigation', *Journal of Money, Credit and Banking* **29**(1), 1–16.
- Johansen, S. (1995), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press.
- Johansen, S. and Swensen, A. R. (1999), 'Testing exact rational expectations in cointegrated vector autoregressive models', *Journal of Econometrics* **93**, 73–91.
- Johansen, S. and Swensen, A. R. (2004), 'More on testing exact rational expectations in cointegrated vector autoregressive models: Restricted constant and linear term', *Econometrics Journal* **7**, 389–397.
- Juselius, K. and Hansen, H. (1995), *CATS in RATS, Cointegration Analysis of Time Series*. User's manual.
- Kara, A. and Nelson, E. (2004), 'International Evidence on the Stability of the Optimizing IS Equation', *Oxford Bulletin of Economics and Statistics* **66**, 687–712. Supplement.
- King, M. (2002), No Money, no Inflation - the Role of Money in the Economy. Bank of England Quarterly Bulletin.
- Linde, J. (2005), Estimating New-Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach. Sveriges Riksbank Working Paper Series, No. 129.
- Ma, A. (2002), 'GMM estimation of the new Phillips curve', *Economics Letters* **76**, 411–417.
- Matheron, J. and Maury, T.-P. (2004), 'Supply-side refinements and the New Keynesian Phillips curve', *Economic Letters* **82**, 391–396.
- Mavroeidis, S. (2004), 'Weak Identification of Forward-looking Models in Monetary Economics', *Oxford Bulletin of Economics and Statistics* **66**, 609–635. Supplement.
- McAdam, P. and Willman, A. (2004), 'Supply, Factor Shares and Inflation Persistence: Re-examining Euro-area New-Keynesian Phillips Curves', *Oxford Bulletin of Economics and Statistics* **66**, 637–670. Supplement.

- McCallum, B. T. and Nelson, E. (1999), ‘An Optimizing IS-LM Specification for Monetary Policy and Business Cycle Analysis’, *Journal of Money Credit and Banking* **31**(3), 298–316.
- Razin, A. (2005), ‘Aggregate Supply, Investment in Capacity, and Potential Output’, *Journal of Money, Credit and Banking* **37**(1), 179–190.
- Roberts, J. M. (1995), ‘New Keynesian Economics and the Phillips Curve’, *Journal of Money, Credit and Banking* **27**(4), 975–984.
- Roberts, J. M. (2005), ‘How Well Does the New Keynesian Sticky-Price Model Fit the Data?’, *Contributions to Macroeconomics* **5**(1), 1–37.
- Romer, C. D. and Romer, D. H. (2004), ‘Choosing the Federal Reserve Chair: Lessons from History’, *Journal of Economic Perspectives* **18**, 129–162.
- Rudd, J. and Whelan, K. (2005), ‘Does the Labor’s Share Drive Inflation?’, *Journal of Money, Credit and Banking* **37**, 297–312.
- Sbordone, A. M. (2002), ‘Prices and unit labor costs: a new test of price stickiness’, *Journal of Monetary Economics* **49**, 265–292.
- Sbordone, A. M. (2005), ‘Do expected future marginal costs drive inflation dynamics?’, *Journal of Monetary Economics* **52**(6), 1183–1197.
- Svensson, L. E. O. (2000), ‘Open-economy inflation targeting’, *Journal of International Economics* **50**, 155–183.
- Tillmann, P. (2005), The New Keynesian Phillips Curve in Europe: does it fit or does it fail? Deutsche Bundesbank Discussion Paper Series 1, No. 4.
- Walsh, C. E. (2003), *Monetary Theory and Policy*, 2 edn, The MIT Press.